Topology Basics

[Channels and Nodes]

a set of nodes $N^* \supseteq$ a set of terminal nodes

a set of channels C

$c \in C = (x, y) \rightarrow x = s_c,\ y = d_c$

$c = (x, y)$: characterized by ...

width W_c or W_{xy}

freq f_c or f_{xy}

latency t_c or $t_{xy} \rightarrow$ length $l_c = v t_c$

v = propagation velocity

bandwidth $b_c = w_c f_c$

$= b$ (if all b_c for $c \in C$ are the same)

For switch node x, $C_x = C_{Ix} \cup C_{Ox}$

where $C_{Ix} = \{c \in C \mid d_c = x\}$

$C_{Ox} = \{c \in C \mid s_c = x\}$

the degree of x : $\delta_x = |C_x|$

$\delta_{Ix} = |C_{Ix}|$: the sum of the in degree

$\delta_{Ox} = |C_{Ox}|$: the sum of the out degree

(if $\forall \delta_x$ are same, just δ).

[Cuts and Bisections]

a cut of networks $C(N_1, N_2)$; $N^* \xrightarrow{\text{Two disjoint}} N_1 \& N_2$

each element in $C(N_1, N_2)$ = a channel with a source/dest. in N_1/N_2

Total BW of the cut : $B(N_1, N_2) = \sum_{c \in C(N_1, N_2)} b_c$

A bisection of a network : $C(N_1, N_2)$ s.t

(partition nodes ~½) $|N_2| \le |N_1| \le |N_2| + 1$

(partition terminal nodes ~½) $|N_2 \cap N| \le |N_1 \cap N| \le |N_2 \cap N| + 1$

- Channel bisection $B_C = \min_{\text{bisection}} |C(N_1, N_2)|$
  (minimum channel count over all bisections)
- Bisection bandwidth $B_B = \min_{\text{bisection}} B(N_1, N_2)$

if uniform channel $\rightarrow B_B = b B_C$ (BW b)

[Paths] = Routes

- a path = ordered set $P = \{c_1, c_2 \cdots c_n\}$
  where $d_{c_i} = s_{c_{i+1}}$

Source of a path $s_P = s_{c_1}$

Dest of " $d_P = d_{c_n}$

- hop count $= |P|$
- "connected" network ? $\exists P_{ij}$ for $\forall$ source $i \in N$ destn $j \in N$
- A set of minimal paths $= R_{xy}$
  the hop count of the minimal paths $= H(x, y)$ between $x \& y$
- $H_{max} = \max_{x, y \in N} H(x, y)$ = The diameter (The largest minimal hop counts over all terminal paths)

eg) $H_{max} \ge \log_{\delta_O} N$; Diameter-lower bound = Tree-type ideal network (symmetric switches)

eg) $H_{max} \ge \log_{\delta/2} N$

- The Average minimum hop count of a network H_{min}

$H_{min} = \frac{1}{N^2} \sum_{x, y \in N} H(x, y)$

- The actual average hop count (Not avg of minimum hop counts)

$H_{avg} \ge H_{min}$

- physical distance of a path

$D(P) = \sum_{c \in P} l_c$

$t(P) = D(P)/v$

[Symmetry]

Vertex symmetry & edge symmetry

[traffic patterns]

message distribution with a traffic matrix Λ

$\lambda_{s,d} \in \Lambda \Rightarrow$ fraction of traffic send from s to d.

- Random traffic
- permutation traffic $d = \pi(s)$
- Bit permutation ; 4-bit source address
  $\{s_3, s_2, s_1, s_0\}$
  Bit-reversed
  $\{s_0, s_1, s_2, s_3\}$
  Bit-complemented
  $\{\neg s_3, \neg s_2, \neg s_1, \neg s_0\}$
  Shuffle
  $\{s_2\ s_1\ s_0\ s_3\}$
  Digit permutation : dest addr. can be calculated from source addr.
  Th... n-digit, radix-k address
  [k-ary n-cube (torus)
  k-ary n-fly (butterfly)

[[Performance]]

[Throughput and Maximum Channel Load]

- Ideal throughput ? w/ perfect flow control & routing
- load on a channel $c = \gamma_c = \left(\frac{\text{BW demand from } c}{\text{BW of the input ports}}\right)$
  $= \left(\frac{\text{traffic demand for } c}{\text{All input ports injects one unit of traffic}}\right)$
- maximum channel load
  $$\gamma_{max} = \max_{c \in C} \gamma_c$$
- Ideal throughput of topology $\boxed{\Theta_{ideal} = \frac{b}{\gamma_{max}}}$
- For uniform traffic → upper/lower bound of γ_{max}!
  ↳ $N/2$ ~~traffic~~ packets must cross B_c bisection channels.
  ∴ Best (uniformly distributed) throughput occurs..
  $$\gamma_{max} \geq \gamma_B = \frac{N}{2B_c}$$
  $\frac{N}{2}$ ⇉ B_c channels

Upper Bound ⇒ $\boxed{\Theta_{ideal} \leq \frac{2bB_c}{N} = \frac{2B_B}{N}}$

(Ring ⇒ Bound is exact)

- $H_{min}N$ = channel demand = # channel traversal
  Ideally distributed

$$\gamma_{c,LB} = \gamma_{max,LB} = \frac{H_{min}N}{C} \quad \text{(eq 3.5)}$$

- If there are $|R_{xy}|$ paths, $1/|R_{xy}|$ is credited and they are 100% overlap

$\gamma_{max,UB}$ = the largest $\gamma_{c,UB}$ over all channels

$$\gamma_{c,UB} = \frac{1}{N}\sum_{x\in N}\sum_{y\in N}\sum_{p\in R_{xy}} \begin{cases} \frac{1}{|R_{xy}|} & \text{if } c \in P \\ 0 & \text{otherwise} \end{cases}$$

$$\gamma_{max,UB} = \max_{c\in C} \gamma_{c,UB} \quad \text{(eq 3.6)}$$

$$\gamma_{max,LB} \leq \gamma_{max} \leq \gamma_{max,UB}$$

(If edge-symmetric topology ~ tori ⇒ exact)

* Computing γ_{max} in general case ?

$\exists_d^{\in N^*}$ (destination) → average distribution of packets destined to d
= vector x_d (length $|x_d| = |C|$)

Flow balancing equation ?

(sum of incoming distribution) − (sum of outgoing channel)
= ⊕ average #packets that the node is sourcing
⊖ " sinking

For uniform ~~XXX~~ traffic ?

All nodes terminal source 1/N, 'd' sinks 1.

(eq)
$$f_{d,i} = \begin{cases} 1/N - 1 & \text{if } d = i \ \&\ d \in N \\ 1/N & \text{if } d \neq i \ \&\ d \in N \\ 0 & \text{otherwise} \end{cases}$$

$|N^*|$ element balance vector f_d

Express the topology using $N^* \times C$ matrix A.

$$A_{n,c} = \begin{cases} +1 & \text{if } s_c = n \\ -1 & \text{if } d_c = n \\ 0 & \text{otherwise} \end{cases}$$

Objective to minimize the maxload on any one channel

minimize $\max_{c\in C} \sum_{d\in N} x_{d,c}$ → optimal max channel load

subject to $Ax_d = f_d$

$x_d \geq 0$ for all $d \in N^*$

Every node is equally likely to send to every node!

Eq 3.6 → weighting final summation by λ_{xy}

(λ_{xy}: probability that x sends to y)

$$\boxed{\gamma_c(\Lambda) = \sum_{x \in N} \sum_{y \in N} \lambda_{xy} \sum_{P \in R_{xy}} \begin{cases} 1/|R_{xy}| & \text{if } c \in P \\ 0 & \text{otherwise} \end{cases}}$$

Eq 3.5 → weighting with probability λ_{xy}

$$\boxed{H_{min}(\Lambda) = \sum_{x \in N} \sum_{y \in N} \lambda_{xy} H(x,y)}$$

- ideal throughput = capacity of the network
- fraction of capacity $= \Theta(\Lambda)/\Theta(U)$

$= \gamma_{max}(U)/\gamma_{max}(\Lambda)$

[latency]

$T = T_h + \frac{L}{b}$

(head lat.) (serialization lat.)

$T_h = T_r + T_w$

T_r (router delay) $= H_{min} t_r$ (single router delay)

T_w (Time of flight) $= D_{min}/v$

$$T_0 = H_{min} t_r + \frac{D_{min}}{v} + \frac{L}{b} \; (+\, T_c)$$

T_0: zero load = no-contention

T_c: time spent waiting for resources (= contention)

[Path diversity]

$(|R_{xy}| > 1)$ → good for robustness of the network.

For permutation traffic ⇒ bottleneck occurs w/o path diversity

As contention ⇑ ⇒ γ_{max} ⇑ ⇒ Θ ⇓

[Packaging cost]

- local wiring: channel width $w \le \frac{W_n}{\delta}$ (W_n: max pin count of a node; δ: degree of node)
- global wiring: $w \le \frac{W_s}{B_c}$ (W_s: # available global wires; B_c: minimum channel bisection)

$\therefore w \le \min\left(\frac{W_n}{\delta}, \frac{W_s}{B_c}\right)$

- max bandwidth $b \le \min\left(\frac{B_n}{\delta}, \frac{B_s}{B_c}\right)$, $B_n = fW_n$, $B_s = fW_s$
- wire length? critical

$f = \min\left(f_0, f_0\left(\frac{l_w}{l_c}\right)^{-2}\right)$

Case study SGI Origin ~~2000~~

	Ring	Cayley
b	35 Gb/s	20 Gb/s
H_{avg}	3/2	7/6
γ_{max}	3/4	7/18
Θ_{ideal}	~46.7 Gb/s	~51.4 Gb/s
T_h	30 ns	23.3 ns
T_s	29.3 ns	51.2 ns
T_0	69.3 ns	74.5 ns